

基于多模态数据的外语教师 TPACK 画像构建 与精准干预研究

潘小燕

广东科技学院, 广东 东莞 523668

DOI: 10.61369/ETR.2026030030

摘要: 随着教育信息化步入“深水区”, 如何实现技术与教学的实质融合已成为关键议题。为探究此问题, 本研究依托 TPACK 理论框架, 采集并分析了外语课堂教学中的多模态行为数据。通过聚类分析, 本研究识别出“技术依赖型”、“教学传统型”、“均衡发展型”与“低活跃型”四类具有代表性的教师群体画像。基于不同群体的特征与困境, 本研究设计了一套差异化的精准干预方案。研究结论表明, 这种数据驱动的动态画像方法, 能够为教师能力的科学评估与个性化培训设计提供依据, 从而推动教育信息化从初期的“形式整合”迈向高质量的“内涵提升”。

关键词: 多模态数据; 教师画像; 精准干预; 教师专业发展

Research on Constructing TPACK Profiles of Foreign Language Teachers Based on Multimodal Data and Precise Interventions

Pan Xiaoyan

Guangdong University of Science and Technology, Dongguan, Guangdong 523668

Abstract: As the advancement of educational informatization enters a "deep-water zone," achieving substantive integration between technology and instruction has become a critical challenge. To address this issue, this study utilizes the TPACK (Technological Pedagogical Content Knowledge) theoretical framework and collects and analyzes multimodal behavioral data from foreign language classrooms. Through cluster analysis, four distinct teacher profiles are identified: "Technology-Dependent," "Teaching-Traditional," "Balanced-Development," and "Low-Activity." Drawing on the characteristics and challenges associated with each profile, the study proposes a set of differentiated, targeted intervention strategies. The findings indicate that this data-driven dynamic profiling approach can provide a scientific basis for assessing teacher competencies and designing personalized training programs, thereby advancing educational informatization from a preliminary stage of "formal integration" to a higher-quality phase of "substantive enhancement."

Keywords: multimodal data; teacher profiling; targeted intervention; teacher professional development

一、背景与目标

在教育数字化转型的宏观背景下, 教师的信息化教学能力已成为衡量教育改革成效的关键指标之一。为系统评估该能力, 本研究援引 Mishra 与 Koehler 提出的 TPACK 理论框架^[1], 该框架超越了将技术、教学法与学科知识简单叠加的视角, 转而强调三者在具体教学情境中深度融合的复杂性与必要性。然而, 传统的教师能力评估多依赖于主观性较强的问卷或观察, 难以捕捉其 TPACK 结构的动态性与实践性全貌。

近年来, 随着教育大数据的积累与分析技术的成熟, 教师在教学过程中留下的多模态行为轨迹——包括但不限于在线平台的交互日志、数字资源的使用偏好以及教学活动的参与深度——为研究者打开了一扇新的窗口。这些颗粒度的行为数据, 为客观、

动态地刻画教师真实的 TPACK 实践水平提供了可能。利用技术对教学过程中的多模态数据进行挖掘与分析, 能够有效提高对教学行为分析的全面性与真实性^[2]。郑晋提出, 传统的教师评价体系难以全面、动态地反映教师教学能力。应将大数据技术、用户画像思想与教育教学理念进行有机融合, 构建可视化的教师教学质量数字画像^[3]。基于教师个体行为特征构建画像, 并据此提供个性化发展支持, 已成为当前教师培训与管理中的关键课题^[4]。

基于上述现实与理论诉求, 本研究以编者所在学校外国语学院超星学习通的后台教学数据为来源, 构建了一个数据驱动的教师信息化教学行为画像模型。整合技术的学科教学知识 (TPACK) 是指导教师将技术有效融入课堂教学的框架, 具备较为成熟的理论体系与测量工具。该框架已在大量研究中被用于评估教师在教学中整合信息技术与课程内容的水平^[5]。具体而言, 本

项目信息: 本文是 2023 年广东科技学院校级科研项目“基于高校混合式教学的外语教师数字素养的提升研究”(GKY-2023KYBW-36) 的研究成果之一。

作者简介: 潘小燕 (1981—), 女, 汉, 广东河源, 副教授, 硕士。研究方向: 英语教育技术、国际传播。

研究采用 K-means 聚类算法对预处理后的教师多模态行为数据进行群体划分，其目的在于发现数据中存在的自然教师群体^[6]。最终，本研究针对不同画像教师群体设计的精准干预方案，遵循教师专业发展的有效性原则，即专业发展应是密集的、持续的、与教学实践紧密相连的^[7]，从而为提升教师信息化教学能力的实践工作提供科学依据与方法论支持。

二、数据预处理

(一) TPACK 理论模型的应用

TPACK 模型的核心理论价值，在于它揭示了成功的技术整合教学并非技术 (T)、教学法 (P) 与学科内容 (C) 三类知识的简单叠加，而是三者交汇融合后所形成的一种全新的、情境化的知识形态^[8]。为将这一理论框架转化为可观测、可度量的行为指标，本研究依据其核心维度，从教师的多模态行为数据中萃取并定义了以下三类特征：

技术应用 (T)：图片、视频、音频、其他格式资源的使用频率。

教学法实践 (P)：创建试卷、发布作业、批阅考试、课堂评分、抢答、发帖 / 回帖等教学活动。

学科内容建设 (C)：创建章节、上传试题、文档使用、在线时长 (内容深度)。

(二) 数据预处理

为保障后续聚类分析的信度与效度，本研究首先对采集到的原始教师行为数据进行了预处理。这一流程主要包括数据清洗与数据标准化两个关键环节。在数据清洗阶段，研究系统性地处理了数据中的缺失值，识别并修正了因系统记录等原因造成的极端异常值，使所有特征具有可比性，满足 K-means 聚类算法对数据分布的基本假设，并对所有数值型特征统一采用了 Z-score 标准化方法。该方法将原始数据转换为均值为 0、标准差为 1 的新数据，其计算公式为：

$$z = \frac{x - \mu}{\sigma}$$

其中，x 为原始值， μ 为特征均值， σ 为特征标准差。

经过上述处理，最终构建了一个高质量、结构化的分析数据集。该数据集的核心特征依据 TPACK 理论框架归纳为以下三类 (部分关键特征示例见表 1)：

表 1 部分关键特征示例

特征类别	具体指标	处理方式
技术应用 (T)	多媒体资源使用指数、平台访问频率、技术工具使用种类	Z-score 标准化、分箱处理
教学法实践 (P)	教学管理活跃度、课堂互动频率、评价行为密度	Z-score 标准化、PCA 降维
学科内容 (C)	章节创建数量、试题更新频率、文档深度修订次数	Z-score 标准化、对数转换

通过上述步骤，原始数据被转化为高质量、可解释的分析数据集，为后续聚类分析与画像构建确定基础。

三、聚类分析方法与过程

在教师信息化行为模式识别的研究中，本研究采用 K-means 聚类算法对预处理后的多维度行为数据进行群体划分。作为无监督学习领域的经典方法，该算法通过迭代优化将样本分配至 K 个簇类，实现组内样本高相似性、组间差异最大化的分类目标。

为确定最优聚类数目，我们综合运用肘部法则与轮廓系数两种评估方法进行验证。具体而言，通过肘部法则观测不同 K 值对应的误差平方和变化规律，发现当 K=4 时 SSE 曲线的下降速率出现显著拐点，暗示继续增加聚类数对模型解释力的提升有限。进一步通过轮廓系数验证，在 K=4 时取得 0.62 的峰值，该数值表明簇内紧密度与簇间分离度达到较优平衡，确认四分方案具有最佳聚类效果。

(一) 聚类算法选择与原理

张瑾等人在分析教师在线学习社区中的会话主题时，采用 K-means 聚类算法的技术路径，对反映教师信息化教学行为的讨论文本进行主题建模，从而有效识别出教师的关注点与认知结构 [9]。本研究还借鉴了王陆的研究，通过综合运用聚类分析、路径分析等方法^[10]，成功从多源数据中识别出不同的教师成长路径，为本研究采用聚类分析来对教师进行 TPACK 画像分类提供了成功的方法论范例。本研究选用 K-means 聚类算法主要基于三方面考量：首先，算法在处理本研究涉及的高维行为数据时展现出良好的计算效率；其次，其简洁的聚类逻辑保证了最终结果具有出色的可解释性，便于后续教学诊断；最重要的是，作为一种经典的无监督学习方法，它特别适用于从尚未标注的教师行为数据中自主发现有意义的群体结构。K-means 算法的优化目标在于通过迭代，将数据划分为 K 个簇，并使簇内样本到其质心的欧氏距离平方和 (SSE) 达到最小化。本研究执行的具体计算流程如下：

初始化：随机选择 K 个初始质心 (本研究采用 K-mean 优化初始化，避免陷入局部最优)。

分配样本：计算每个样本到各质心的距离，将其分配到最近的簇。

更新质心：重新计算每个簇的均值作为新质心。

迭代优化：重复步骤 2-3，直到质心变化小于阈值 (本研究设定为 $\sqrt{10^{-4}}$) 或达到最大迭代次数 (100 次)。

(二) 最佳聚类数确定

为确定数据的最优簇数 K，结合以下两种方法进行验证：

1. 肘部法则 (Elbow Method)

原理：通过计算不同 \sqrt{K} 值下的 SSE，寻找 SSE 下降速率显著减缓的“拐点”。

实现过程：

· 遍历 $K=1$ 至 $K=10$ ，计算每个 K 对应的 SSE (公式如下)：

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} \|x - \mu_i\|^2$$

其中 C_i 为第 i 个簇， μ_i 为簇中心。

· 绘制 SSE 随 K 变化的曲线 (图 1)，当 $K=4$ 时曲线

斜率骤减，形成明显“肘部”，表明增加簇数对模型提升有限。

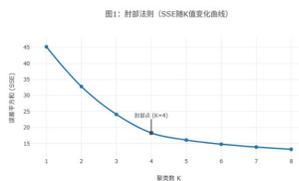


图1：肘部法则（SSE随K值变化曲线）

2. 轮廓系数 (Silhouette Score)

原理：衡量样本与同簇其他样本的紧密度 $a(i)$ 和与最近邻簇样本的分离度 $b(i)$ ，取值范围为 $[-1, 1]$ ，值越接近1表示聚类效果越好。

计算步骤：

对每个样本 i ，计算其轮廓系数：

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$$s(i) = \max\{a(i), b(i)\} - a(i)$$

对所有样本的 $s(i)$ 取均值，得到整体轮廓系数。

测试不同 K 值 ($K=2-8$)，当 $K=4$ 时轮廓系数达到峰值0.62 (图2) 表明簇内紧密且簇间分离度高。

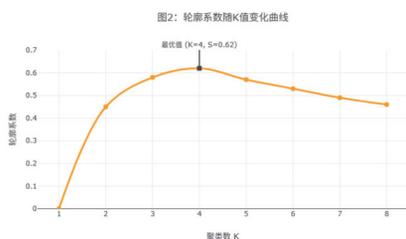


图2：轮廓系数随K值变化曲线

(三) 数据输入与参数设置

· 输入数据：预处理后的数据集，包含技术应用 (T)、教学法实践 (P)、学科内容 (C) 三类共12个特征。所有特征均已通过 Z-score 标准化，消除了量纲影响，为聚类分析提供了可比且无偏的数据基础。

· 距离度量：采用欧氏距离，公式为：

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

由于输入数据已标准化，各特征对距离计算的贡献是均衡的。

(四) 聚类结果验证：簇内特征分析

通过计算各簇中心在 T、P、C 三个维度的均值 (表2)，定义四类教师群体：

表2 T、P、C 三个维度的均值

聚类标签	技术应用 (T)	教学法实践 (P)	学科内容 (C)	样本占比
技术依赖型 (C1)	高 (1.32)	低 (0.21)	低 (0.18)	22%
教学传统型 (C2)	低 (0.15)	高 (1.45)	中 (0.76)	34%

聚类标签	技术应用 (T)	教学法实践 (P)	学科内容 (C)	样本占比
均衡发展型 (C3)	中 (0.89)	中 (0.93)	高 (1.28)	28%
低活跃型 (C4)	低 (0.08)	低 (0.12)	低 (0.09)	16%

(五) 鲁棒性检验

为验证结果的稳定性，采用以下方法：

多次随机初始化：重复运行 K-means 算法10次，簇中心差异小于5%，表明结果稳定。

外部指标验证：与人工标注的教师能力等级 (低、中、高) 进行交叉检验，聚类结果与人工评价的一致性系数 (Cohen's Kappa) 为0.71 ($p < 0.01$)，表明聚类具有较高可信度。

(六) 方法局限性

聚类结果受限于数据质量与特征工程的设计，未来需结合质性数据 (如教学反思日志) 增强解释性。且当前分析基于静态截面数据，未考虑教师行为的时序演化，后续可引入时间序列聚类方法。此外，本研究为提升聚类效果的统一性，对所有特征统一采用了 Z-score 标准化。未来研究可尝试对比不同标准化或加权策略对教师画像结果的影响，以探索更精细化的特征工程方案。

四、结果分析

(一) 聚类结果

从方法选择、参数优化到结果验证，最后通过分析各簇中心点特征，系统化将教师群体划分为4类信息化行为模式：

表3 教师群体类别分布与占比

聚类标签	技术应用 (T)	教学法实践 (P)	学科内容 (C)	样本占比
技术依赖型 (C1)	高 (1.32)	低 (0.21)	低 (0.18)	22%
教学传统型 (C2)	低 (0.15)	高 (1.45)	中 (0.76)	34%
均衡发展型 (C3)	中 (0.89)	中 (0.93)	高 (1.28)	28%
低活跃型 (C4)	低 (0.08)	低 (0.12)	低 (0.09)	16%

(二) 教师画像分类

胡小勇等提出，面对新时代教师队伍建设的的要求，教研向个性化、精准化转型势在必行^[10]。通过以上聚类算法结果，教师群体划分为4类信息化行为模式具体如下：

1. 技术依赖型

“技术依赖型”教师群体呈现出鲜明的行为特征：他们对各类技术工具表现出极高的使用热情与频率，然而，其技术应用与深层教学目标及学科内容本质之间的融合度明显不足。这类教师善于利用丰富的多媒体资源 (如视频、交互式动画) 来营造生动的课堂氛围，但在实践中，这种技术应用往往未能有效服务于教学内容，不自觉地陷入了“技术展示优先”的形式化误区。

以英语听力教学为例，此类教师的设计思路可能是率先引入某款语音识别软件来呈现材料，而非从诊断学生特定的听力障碍出发来规划教学路径。这种“工具先行”的模式，常导致技术应用与教学核心目标产生错位。数据进一步揭示了这一矛盾，该群体在平台的总体在线时长普遍高于均值，但其投入到核心教学活动（如作业批改）的精力却相对有限。典型的表现是，教师可能投入大量时间精心制作了视觉效果突出的PPT课件，但由于未能将课件内容与学生的认知难点、学习节奏进行精准对接，最终的教学效果提升有限，技术的投入产出比不尽如人意。

2. 教学传统型

与“技术依赖型”教师形成鲜明对比，“教学传统型”教师群体展现出对经典教学范式的路径依赖。数据显示，他们在作业发布、试卷批阅等核心教学环节上保持着高度的活跃性，但其技术工具的使用广度与深度均显著低于样本平均水平。

该群体的信息化行为呈现出明显的“工具保守性”特征，其数字活动大多局限于课程章节构建、文档资料上传等基础性、非交互式的功能范畴。这反映出他们将技术平台更多地视为一种替代传统黑板与讲义的数字“陈列柜”，而非能够重塑教学流程与互动模式的创新工具。因此，其TPACK结构中的技术知识（TK）与其他维度的融合程度明显不足，技术整合能力整体薄弱，尚未能有效利用技术来赋能与深化其固有的教学优势。

3. 均衡发展型（Cluster3）

均衡发展型教师群体在技术应用、教学法实践与学科内容建设三个维度均表现活跃，其TPACK整合能力处于领先水平。

作为教师队伍中的“引领型”群体，均衡发展型教师已展现出技术应用、教学法创新与学科内容深度融合的显著特质。这类教师不仅能够熟练运用各类信息技术工具，更善于从教学逻辑出发，将技术有机嵌入学科教学的整体架构之中，形成良性互动的教学闭环。其TPACK结构呈现出动态平衡、相互促进的成熟特征。

4. 低活跃型（Cluster4）

低活跃型教师群体在技术应用、教学法实践与学科内容建设三方面均表现薄弱，其信息化教学参与度显著低于其他群体。低活跃型教师在技术接受度、教学法更新与学科内容建设等多个维度均存在明显的发展滞后，其信息化教学实践频率与质量显著落后于其他群体。这种发展困境往往源自多重因素的叠加影响：既包括技术信心的缺乏、专业发展支持的不足，也可能涉及教学理念更新滞后等深层次问题。

（三）精准干预策略

1. 技术依赖型（Cluster1）

基于TPACK理论框架，针对技术依赖型的教师干预需聚焦技术、教学法与学科内容的深度融合，从培训设计与资源优化两方面提出系统性解决方案。培训应围绕“目标导向的技术整合”展开。设计基于学科的教学场景，引导教师将技术工具与具体教学目标匹配。例如，在英语写作教学中，若目标是“提升学生议论文逻辑性”，可引导教师优先选择思维导图工具辅助论点梳理，而非盲目使用视频资源。

为教师提供可借鉴的整合范例，缓解其“有技术、无策略”的困境，如模态案例库，按学科分类，收录优秀课例视频、课件源码及设计说明。例如，收录“基于虚拟现实（VR）的西班牙语文化沉浸课”完整教案，且标注其TPACK整合点，如VR技术如何服务于跨文化交际能力目标。以教研室为单位，开展“技术整合示范课”观摩与评析，鼓励教师互评技术应用的有效性。利用教学平台的行为日志，监测教师课件互动率、学生任务完成度等指标，动态评估干预效果。设立“TPACK整合创新奖”，表彰在技术与学科融合中表现突出的教师，形成正向引导。

2. 教学传统型（Cluster2）

针对教学传统型群体的干预需聚焦技术工具的低门槛引入与教学法升级。教学传统型教师的技术焦虑较高，需通过低门槛工具降低学习成本。以具体教学任务为切入点，演示技术工具如何提升效率。例如，在英语阅读教学中，展示“讯飞语记”如何通过语音转文字功能快速记录学生讨论内容。

引导教师将技术融入传统教学环节，培训教师利用微课视频实现课前自学，腾出课堂时间进行深度讨论。引入课堂即时反馈工具，将技术应用于学生过程性评价，替代单一纸质测试。在教研活动中强制要求使用1-2项技术工具，培养基础使用习惯。组织“技术+教学法”融合示范课，展示技术如何优化传统教学环节。将技术应用纳入教师职业发展体系，与教师能力考核挂钩。

3. 均衡发展型（Cluster3）

通过开放资源与生态化支持，推动其成为信息化教学的引领者。引入前沿技术工具，拓展教学创新边界。培训使用“Deepseek教学助手”生成个性化学习路径，有条件的情况下教授ChatExcel基础与数据可视化工具，支持教师自主分析学生行为数据，优化教学决策。

搭建经验共享与协同创作平台，通过社交媒体或论坛，按学科技术主题划分社群如“AI+外语教学”，定期举办线上案例分享会。鼓励教师上传技术整合课件与微课视频，采用开放共享，形成可持续的资源再生机制。组织均衡发展型教师与低活跃型教师结对，通过“导师制”推动全校TPACK能力均衡化。建立教师研修机制，通过课题研究、学术工作坊等方式推动其从实践者向教学引领者转型。

4. 低活跃型（Cluster4）的精准干预策略

针对低活跃型教师群体的干预需采取策略，通过基础能力重建与激励机制设计，逐步提升其信息化教学意愿与能力。

对该群体的干预需采取先诊断后支持的策略。首先，通过深度访谈与课堂观察，识别每位教师的具体发展障碍，形成个性化诊断报告；其次，设计技能提升路径，将复杂的技术整合任务分解为可达成的基础模块，通过小步快走的方式重建教学信心；再次，组建由骨干教师牵头的小组，建立常态化帮带机制；从零起点切入，消除技术使用恐惧，为每位低活跃型教师配备1名均衡发展型教师作为导师，提供“手把手”操作指导。同时开发5分钟以内的微课系列，支持教师按需碎片化学习。再通过任务驱动与正向反馈激发参与意愿，设计从易到难的任务链。

五、结论

本研究实证表明,基于TPACK理论框架,通过聚类分析方法所构建的“信息化教学行为画像”,能够超越传统的主观评价,实现对教师群体信息化教学行为特征与能力短板的精准诊断与可视化呈现。这一方法将抽象的教师能力转化为具象的行为模式图谱,为教育管理者设计差异化的教师培训体系、实施精准的资源推送提供了坚实的学理依据与数据支持。

通过对教师行为数据进行持续的动态监测与周期性的聚类分析,可以敏锐捕捉教师群体行为模式的演变趋势与干预策略的实际效果。基于这些动态反馈,教育管理者能够对培训方案与资源分配策略进行迭代优化,从而实现从静态的、一次性的“群体画像”到动态的、发展性的“成长档案”的转变。这一闭环机制有望系统性地、可持续地推动整个教师群体信息化教学能力从“形式整合”向“质量提升”的深层演进。

参考文献:

- [1]MISHRA P, KOEHLER M J. Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge[J]. Teachers College Record, 2006, 108(6): 1017-1054.
- [2]张乐乐, 顾小清. 多模态数据支持的课堂教学行为分析模型与实践框架[J]. 开放教育研究, 2022, 28(06): 101-110.
- [3]郑晋. 基于大数据的高校教师教学质量画像的研究与实现[J]. 信息系统工程, 2023, (06): 173-176.
- [4]燕学敏. 精准培训视域下教师专业自画像指标体系的建构[J]. 教学与管理, 2025, (15): 46-51.
- [5]Jain, A. K. Data clustering: 50 years beyond K-means[J]. Pattern Recognition Letters, 2010, 31(8): 651 - 666.
- [6]王晶心, 王胜清, 陈文广. 基于TPACK的高校教师混合式教学胜任力模型研究[J]. 中国远程教育, 2022, (08): 26-34.
- [7]DESIMONE L M. Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures[J]. Educational Researcher, 2009, 38(3): 181-199.
- [8]阮全友, 杨玉芹. 整合技术的学科和教学知识框架的发展: 从TPACK、TSACK到TMACK[J]. 中国远程教育, 2014, (11): 20-26+96.
- [9]张瑾, 张爽楠, 叶海智, 等. 教师在线学习社区中会话主题发现及演化分析[J]. 远程教育杂志, 2021, 39(02): 85-94.
- [10]王陆, 彭功, 马如霞, 等. 大数据知识发现的教师成长行为路径[J]. 电化教育研究, 2019, 40(01): 95-103.
- [11]胡小勇, 林梓柔. 精准教研视域下的教师画像研究[J]. 电化教育研究, 2019, 40(07): 84-91.